



DOI: 10.5862/JCSTCS.229.8

УДК 025.4, 004

*В.Н. Корелин, И.С. Блеканов, С.Л. Сергеев*

## **ПРИМЕНЕНИЕ МОДИФИЦИРОВАННОГО АЛГОРИТМА LSH ДЛЯ КЛАСТЕРИЗАЦИИ ВНЕШНЕГО ОКРУЖЕНИЯ ВЕБ-ПРОСТРАНСТВА УНИВЕРСИТЕТОВ\***

*V.N. Korelin, I.S. Blekanov, S.L. Sergeev*

### **CLUSTERING OF THE EXTERNAL WEB ENVIRONMENT OF UNIVERSITIES USING A MODIFIED LSH ALGORITHM**

Проведен кластерный анализ внешних ресурсов сайтов крупных университетов. В качестве исследуемых объектов выбраны сайты университетов России, США и Великобритании, занимающие в своих регионах ведущие позиции в вебометрическом рейтинге. Цель работы — в выявлении для каждого сайта университета групп внешних веб-ресурсов с одинаковым родом деятельности. Проведен анализ найденных групп: определена степень влияния количества и размеров этих групп на вебометрический рейтинг сайтов университетов. Разработан алгоритм кластеризации, основанный на вероятностном методе понижения размерности многомерных данных (Locality-Sensitive Hashing — LSH). Поставлен эксперимент, в котором на тестовых данных показано, что алгоритм позволяет с высокой скоростью и допустимой точностью проводить кластеризацию большого объема данных. Приведены основные результаты исследования.

ВЕБОМЕТРИКА; ВЕБ-САЙТЫ УНИВЕРСИТЕТОВ; КЛАСТЕРНЫЙ АНАЛИЗ; LOCALITY-SENSITIVE HASHING; MIN HASHING; КЛАСТЕРИЗАЦИЯ ВНЕШНИХ ВЕБ-РЕСУРСОВ; АНАЛИЗ ГИПЕРССЫЛОК.

The paper is dedicated to cluster analysis of external web sites of large universities (web sites that refer to universities and web sites that are referred by universities). Web sites in Russia, the USA and the UK that have highest webometric ranking in their region were chosen as the subject of the study. The goal of the research is to identify a group of sites for each university that have the same kind of activity. The found clusters have been analyzed to determine the impact of group size and the number of groups on webometric ranking of university sites. To achieve the goal of the research, the authors developed a clustering algorithm based on the probabilistic method of reducing the dimension of multidimensional data (Locality-Sensitive Hashing, or LSH). An experiment that was conducted using the test data showed that the developed algorithm has good clustering quality and fast speed performance during massive dataset mining. The main results of the research are presented.

WEBOMETRICS; EXTERNAL WEB SITES OF UNIVERSITIES; CLUSTERING; LOCALITY-SENSITIVE HASHING; MIN HASHING; EXTERNAL WEB SITES CLUSTERING; HYPERLINKS ANALYSIS.

Для большинства крупных организаций немаловажное значение имеет их рейтинг, который рассчитывается в зависимости от

параметров, связанных с их родом деятельности. В частности, на общий рейтинг вуза большое влияние оказывает его вебометри-

\* Данная статья является расширенной версией работы «Hierarchical Clustering of Large Text Datasets Using Locality-Sensitive Hashing» (авторы Vasilii Korelin, Ivan Blekanov), представленной на конференции International Workshop on Applications in Information Technology (IWAIT-2015) и опубликованной в материалах конференции: Pyshkin E.,

and Klyuev V. (Eds). Proceedings of the International Workshop on Applications in Information Technology (IWAIT-2015), The University of Aizu Press, 2015. Электронная версия доступна на сайте <<http://kspt.ftk.spbstu.ru/media/files/2015/iwait-2015/proceedings/iwait-2015-e-proceedings-release.pdf>>

ческий рейтинг. Известно, что одним из главных показателей, влияющих на веб-метрический рейтинг любой организации, в т. ч. и университета, является количество внешних ресурсов [1], ссылающихся на сайт университета. Но кроме количества ссылающихся ресурсов также важно понять качество этих ресурсов, их природу, определить к какой области относится тот или иной внешний ресурс. Данные веб-ресурсы образуют группы сайтов с одинаковым родом деятельности. Таким образом, возникает задача выявления этих групп — кластеризации. Требуется определить степень влияния количества и размеров найденных групп на вебметрический рейтинг сайтов университетов. По найденным кластерам можно определить, с какими группами внешних ресурсов следует сайтам университетов выстраивать гиперссылочные взаимосвязи для повышения цитируемости.

Объектом данного исследования являются университетские сайты, имеющие чрезвычайно большие размеры (например, сайт СПбГУ содержит более 50 тыс. внутренних веб-страниц и около 5 млн гиперссылок) [2, 3]. Окружение таких сайтов может составлять десятки-сотни тысяч страниц. Стандартными методами кластеризации такого объема внешних веб-ресурсов не обойтись, т. к. при работе с большими коллекциями документов многие из данных методов показывают крайне неудовлетворительные результаты в плане производительности [4] (например, метод Single Linkage позволяет создавать кластеры произвольной формы, однако имеет высокую трудоемкость —  $O(n^2)$ , где  $n$  — число документов). Для того чтобы избавиться от «проклятия» размерности применяется вероятностный метод понижения размерности многомерных данных Locality-Sensitive Hashing — LSH, основная идея которого состоит в подборе хэш-функций для некоторых измерений для того, чтобы похожие объекты попадали в одну корзину [5].

#### **Уменьшение размерности для анализа больших коллекций текстовых документов**

Для преобразования текстовых документов в числовые множества в данной работе

использовался метод Shingling [6], который разбивает каждый документ на небольшие множества по  $k$  слов в каждом. Далее для сравнения документов применялась технология Min Hashing, позволяющая быстро сравнивать множества, содержащие большое число элементов.

*Мера Жаккара.* В работе для определения схожести двух текстов применяется мера Жаккара, определяющая схожесть множеств отношением числа элементов, входящих в пересечение двух множеств к числу элементов, входящих в объединение этих множеств [7]:

$$\text{Sim}(C_1, C_2) = \frac{|C_1 \cap C_2|}{|C_1 \cup C_2|}.$$

Так как каждый документ преобразуется во множества, состоящие из  $k$  слов, то набор документов можно представить в виде сильно разреженной булевой матрицы, где столбцы представляют собой документы, а строки — элементы универсального множества (например, множество элементов, где каждый элемент представляет собой множество  $k$  слов). Элемент такой матрицы равен единице, если документ (столбец) содержит данное множество  $k$  слов. В противном случае элемент матрицы равен нулю.

*Minhashing.* Для большой коллекции документов булева матрица будет сильно разрежена. Соответственно, матрица, описывающая такую коллекцию, займет много места в памяти. Кроме того, дальнейшая ее обработка также займет большое количество времени. Для того чтобы решить возникшие проблемы, данную матрицу преобразуем в матрицу, хранящую определенное количество хэш-функций и информацию о сходстве между похожими документами.

*Minhash-функция  $h(C)$*  — это номер первой строки для столбца  $C$  в булевой матрице, где строки перемешаны случайным образом [8].

Как видим, число случайных перестановок задает число Minhash-функций. Например, можно использовать сто случайных перестановок для создания ста сигнатур для каждого столбца матрицы.

Сигнатуры могут быть записаны в другой матрице сигнатур (Signature matrix), чьи



колонки представляют собой документы, а строки — Minhash-значения.

Чем больше хэш-функций, тем выше вероятность того, что  $\text{Sim}(C_1, C_2) = \text{Sim}(M[C_1], M[C_2])$ , где  $C_i$  — столбец булевой матрицы,  $M[C_i]$  — столбец матрицы сигнатур. Это важное свойство позволяет преобразовывать булевы матрицы с большим количеством строк в небольшие матрицы сигнатур, сохраняя сходство похожих множеств. Соответственно, похожесть двух столбцов определяется долей строк, в которых они равны.

Таким образом, каждый документ может быть представлен в виде вектора, число элементов которого равно количеству Minhash-функций.

**Locality-Sensitive Hashing.** Основная идея: сгенерировать из большого множества документов маленькие списки пар документов, чья схожесть должна быть посчитана. Для сравнения двух документов (столбцов) устанавливается порог  $t$  ( $t < 1$ ). Пара документов считается похожей только в том случае, если доля одинаковых значений в матрице сигнатур больше  $t$ . Для матрицы сигнатур необходимо несколько раз вычислить хэш-значения и поместить документы с одинаковым значением в одну корзину (*bucket*). Документы, которые хоть раз попали в одну корзину, будут рассмотрены как кандидаты на сравнение [9].

Для множественного подсчета хэш-функций столбцов необходимо разбить матрицу сигнатур  $M$  на  $b$  частей по  $r$  строк в каждой. Для каждого  $b$  подсчитать хэш-значения столбцов и поместить столбцы с равным значением в одну корзину. Кандидатами на сравнение будут те столбцы, которые хоть раз попали в одну корзину. Для правильной работы алгоритма необходимо настроить  $r$  и  $b$  таким образом, чтобы похожие документы попадали в одну корзину, а непохожие — в разные.

#### Иерархическая кластеризация с использованием LSH

Данный алгоритм кластеризации использует хэш-таблицы, сформированные в результате LSH [10]. Данный алгоритм кластеризации с высокой долей вероятности создает такие же кластеры, как и в методе

Single linkage. Ниже представлено детальное описание алгоритма.

*Предварительные условия:*

$t < 1$  — порог, задающий минимальную схожесть документов;

$r = 1$  — начальное значение строк матрицы сигнатур в каждой группе;

$r_{\min}$  — минимальное значение строк матрицы сигнатур в каждой группе;

$\Delta$  — коэффициент уменьшения параметра  $r$ .

Каждый документ представляет собой отдельный кластер.

**Шаг 1.** Для каждой группы  $b$  в каждом столбце вычислить хэш-функции и сохранить столбцы, у которых хэш-значения хотя бы раз попали в одну корзину, при условии, что в корзине должны находиться столбцы, принадлежащие разным кластерам. Если, например, какие-то два столбца принадлежат одному кластеру, то один из случайно выбранных столбцов удаляется из корзины.

**Шаг 2.** Для каждого из столбцов, входящих в одну корзину, отобрать пары кластеров, расстояние между которыми больше  $t$ .

**Шаг 3.** Пары кластеров, соответствующие парам, полученным на шаге 2, объединяются в один.

**Шаг 4.** Если  $r \leq r_{\min}$ , то алгоритм прекращает работу. Иначе, переход на шаг 5.

**Шаг 5.**  $r = r - \Delta$ . Переход на шаг 1.

*Оценка качества кластеризации на основе LSH.* Для оценки качества модифицированного метода агломеративной кластеризации с использованием LSH в данной работе использовалась тестовая коллекция текстовых документов Reuters-21578, содержащая 21 578 документов. На данной коллекции оценивались два метода иерархической кластеризации: стандартный алгоритм иерархической кластеризации Single-Link и модифицированный алгоритм Single-Link, основанный на применении LSH. В качестве основных метрик качества в работе использовались точность ( $R$ ), полнота ( $P$ ), аккуратность ( $Acc$ ) и  $F$ -мера.

В табл. 1 приведены результаты оценки качества иерархической кластеризации для каждого из алгоритмов, а также время ра-

Таблица 1

Результаты оценки качества кластеризации на коллекции Reuters

Алгоритм	Количество документов	<i>Acc</i> , %	<i>R</i> , %	<i>P</i> , %	<i>F</i> -мера, %	Время работы
Single-Link	1000	75	79	60	68	4 с
	10 000	7	82	62	71	410 с
	20 000	—	—	—	—	> 1 ч
Single-Link + LSH	1000	72	81	66	73	3 с
	10 000	72	80	69	74	41 с
	20 000	78	83	64	72	90 с

боты каждого алгоритма при кластеризации 1000, 10 000 и 20 000 документов.

Как видно из таблицы, аккуратность и полнота метода Single-Link + LSH близки по значениям к аккуратности и полноте метода Single-Link. Точность разработанного метода иногда превосходит точность метода Single-Link. *F*-меры у обоих методов примерно одинаковые.

На рис. 1 приведена зависимость продолжительности работы алгоритмов кластеризации от числа входных документов.

Графики показывают, что алгоритм кластеризации с использованием LSH работает значительно быстрее, чем алгоритм без использования LSH. С увеличением числа

документов время работы алгоритма с использованием LSH линейно растёт.

#### Эксперимент. Исследование сайтов университетов методом кластеризации на основе LSH

**Постановка эксперимента.** В эксперименте требовалось для заданного списка сайтов университетов России, США и Великобритании, занимающих по своим регионам ведущие позиции в вебметрическом рейтинге [1], с помощью специализированного поискового робота [11] и базы данных Majestic [12] получить списки и содержимое всех внешних веб-страниц, которые их цитируют. А также с помощью

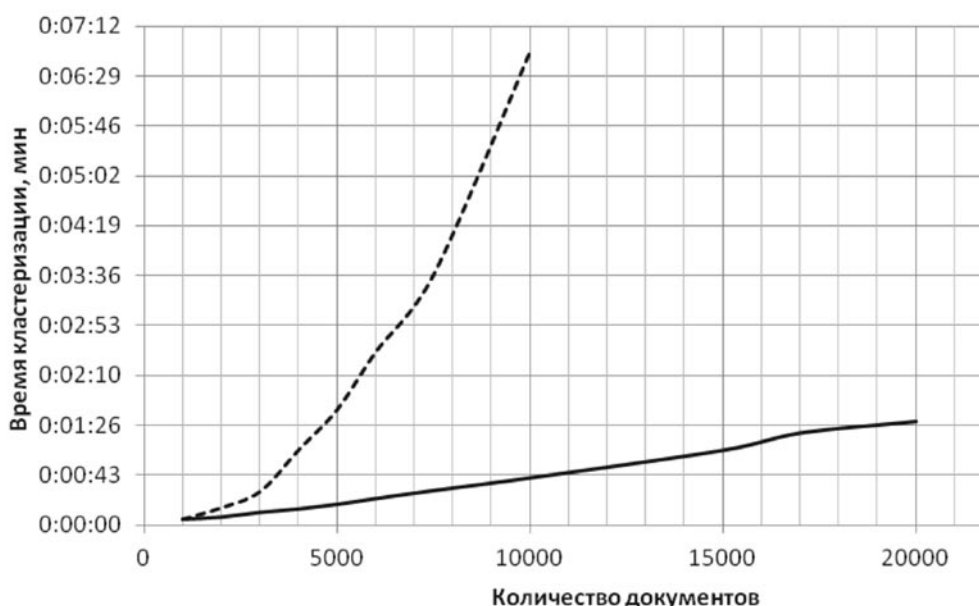


Рис. 1. Время работы алгоритмов кластеризации  
(- - -) Single-Link; (—) Single-Link + LSH



Таблица 2

## Количество внешних ресурсов, окружающих университеты

URL-адрес вуза	Позиция в рейтинге Webometrics	Внешние ссылки на сайте вуза	Внешние домены на сайте вуза	Количество цитирующих сайт вуза веб-страниц	Количество цитирующих сайт вуза доменов
spbu.ru	539	3 280	469	1 599 059	16 028
msu.ru	129	817	262	7 039 127	39 416
nsu.ru	616	5 747	921	926 004	15 186
harvard.edu	1	918	163	75 994 723	319 445
stanford.edu	2	175	79	29 551 130	311 148
mit.edu	3	9	6	41 271 678	324 989
ox.ac.uk	16	8 075	1 631	8 920 524	117 959
cam.ac.uk	15	3 385	1 174	12 084 107	120 796
ucl.ac.uk	24	25 218	5 683	4 733 566	66 035

разработанного авторами метода агломеративной кластеризации на основе алгоритма LSH выявить целевые группы найденных внешних веб-ресурсов с одинаковым родом деятельности и установить степень их влияния на вебометрический рейтинг сайтов исследуемых вузов.

Для исследования были выбраны следующие сайты вузов:

Московский государственный университет имени М.В. Ломоносова (msu.ru), Россия;

Санкт-Петербургский государственный университет (spbu.ru), Россия;

Новосибирский государственный университет (nsu.ru), Россия;

Массачусетский технологический институт (mit.edu), США;

Гарвардский университет (harvard.edu), США;

Стэнфордский университет (stanford.edu), США;

Кембриджский университет (cam.ac.uk), Великобритания;

Оксфордский университет (ox.ac.uk), Великобритания;

Университетский колледж Лондона (ucl.ac.uk), Великобритания.

**Результаты сбора.** В итоге сбора данных

для анализа внешних веб-ресурсов сайтов исследуемых университетов были получены следующие результаты, представленные в табл. 2.

Таблица показывает, что университеты США, занимающие первые три позиции вебометрического рейтинга, имеют значительно больше внешних ресурсов, ссылающихся на них, чем ведущие университеты России и Великобритании. Ведущие университеты Великобритании также опережают российские по общему количеству ссылок и доменов, ссылающихся на них. Количество внешних ресурсов напрямую влияет на такой вебометрический индикатор, как Impact ресурса в Вебе.

На показатель видимости ресурса в Вебе, помимо числа внешних ссылающихся ресурсов, также влияет и их качество. Для кластеризации внешних веб-ресурсов были выбраны только англоязычные и русскоязычные страницы Веба. В табл. 3 приведено количество доменов для каждого университета, к которым применялся метод агломеративной кластеризации с использованием LSH.

Ниже приведены результаты кластеризации внешних ресурсов. Для определения их тематики извлекались наиболее часто

Таблица 3

Количество внешних ресурсов для кластеризации

Университет	Внешние домены на сайте вуза		Цитирующие сайт вуза домены	
	на русском	на английском	на русском	на английском
spbu.ru	342	81	7 373	3 929
msu.ru	177	62	3 343	1 008
nsu.ru	535	247	2 508	1 428
harvard.edu	4	151	1 545	16 632
stanford.edu	5	65	409	6 088
mit.edu	3	3	1 625	24 930
ox.ac.uk	29	1 399	126	3 187
cam.ac.uk	91	4 668	671	9 196
ucl.ac.uk	14	1 054	162	3 112

встречающиеся множества слов (shingles), которые присутствовали в кластере. В результате анализа были выделены основные часто встречающиеся группы: университеты, научные сообщества, поисковые системы, социальные сети (включая различные блоги, форумы и т. д.), медиасфера, новостные порталы и сайты других организаций.

Таким образом, определялась численность каждой группы для университета.

**Результаты кластеризации внешних доменов на сайте вуза.** Внешние домены — это домены, на которые ссылаются исследуемые сайты университетов. Ввиду ограниченного числа внешних ссылок на сайтах университетов США, рассматривались только

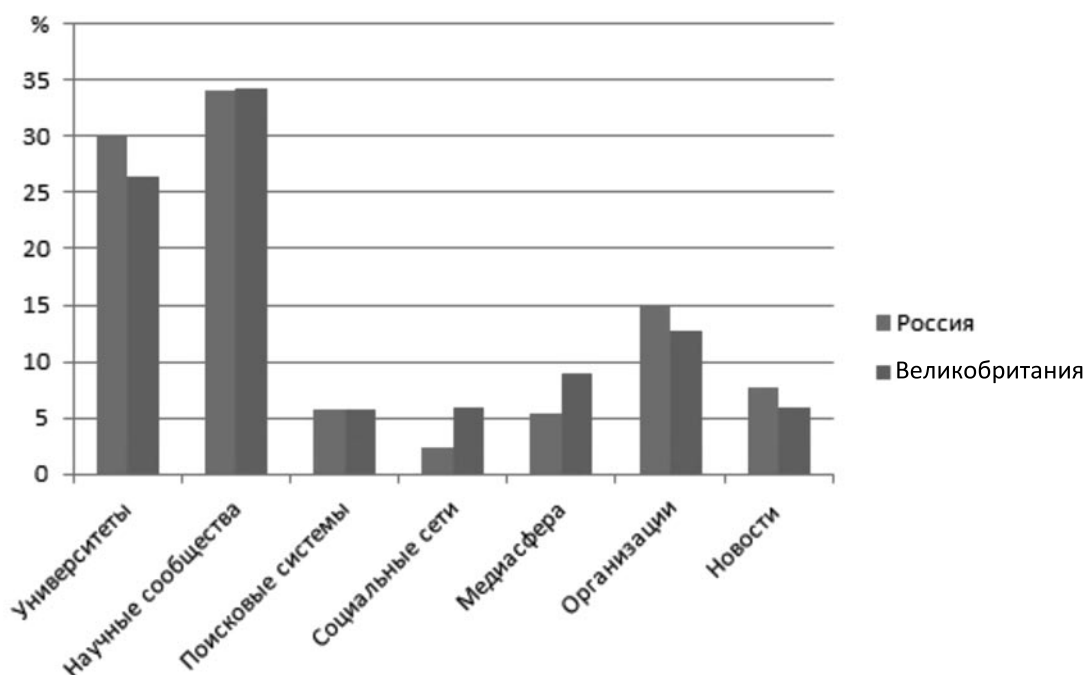


Рис. 2. Сравнительная гистограмма университетов России и Великобритании

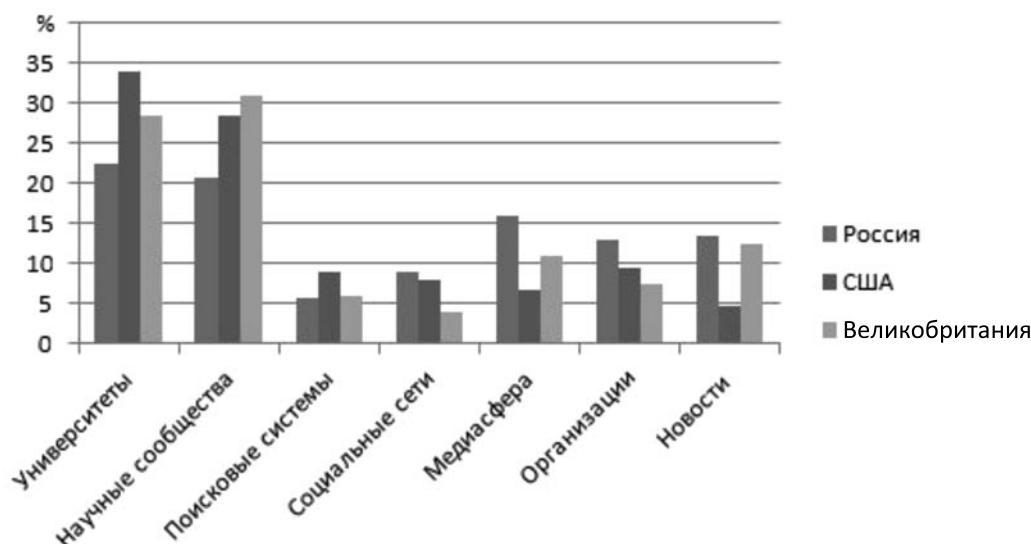


Рис. 3. Сравнительная гистограмма ресурсов, ссылающихся на университеты России, США и Великобритании

внешние ресурсы университетов России и Великобритании. Были получены кластеры для каждого из исследуемых университетов. Ниже представлена сравнительная гистограмма внешнего окружения университетов России и Великобритании (рис. 2).

Рисунок показывает, что внешние ресурсы, на которые ссылаются сайты университетов, практически совпадают. Университеты и научные сообщества составляют значительную долю внешних ресурсов, на которые ссылаются сайты университетов России и Великобритании.

**Результаты кластеризации цитирующих сайт вуза доменов.** Цитирующие веб-ресурсы – это сайты, сгруппированные по доменам, которые ссылаются на исследуемые сайты университетов. Были получены кластеры для каждого из исследуемых университетов. На рис. 3 приведена сравнительная гистограмма внешнего окружения университетов, представляющих Россию, США и Великобританию.

Анализ сайтов, ссылающихся на сайты вузов, показал, что университеты из одной страны имеют схожее внешнее окружение во многих областях. У всех университетов преобладают такие группы, как «Университеты» и «Научные сообщества». У университетов США и Великобритании эта доля выше, чем у российских. Соответственно,

в других областях доля российских университетов больше.

Можно сделать вывод о том, что чем выше доля университетов и научных сообществ среди внешних ресурсов университета, тем выше его вебметрический показатель Impact. Этот показатель зависит от количества и качества внешних ресурсов, ссылающихся на сайт. Внешние группы, содержащие университеты и научные сообщества, имеют значительно больший вес для вебметрического рейтинга, чем другие группы внешних ресурсов.

Разработан алгоритм иерархической кластеризации с использованием вероятностного метода понижения размерности многомерных данных Locality-Sensitive Hashing, оптимально подходящий для кластеризации больших коллекций документов (massive datasets). Данный алгоритм апробирован на тестовой коллекции текстовых документов Reuters. На тестовой коллекции алгоритм показал приемлемую точность (ассигасу) и  $F$ -measure в сравнении с классическими методами иерархической кластеризации. Метод кластеризации с использованием LSH значительно превосходит по скорости работы классические методы иерархической кластеризации.

Разработанный алгоритм использовался для анализа веб-пространства нескольких университетов с их окружением.

Получены данные о внешнем окружении веб-пространства университетов Рос-

сии и Великобритании и о ресурсах, ссылающихся на сайты университетов России, США и Великобритании.

Работа выполнена при финансовой поддержке РФФИ, грант № 15-01-06105.

## СПИСОК ЛИТЕРАТУРЫ

1. Ranking Web of Universities [электронный ресурс]/ URL: [www.webometrics.info](http://www.webometrics.info) (дата обращения: 10.11.2015).

2. Блеканов И.С., Москалец Р.Ю. Теоретико-графовые характеристики в вебметрических исследованиях внутренней топологии крупных сегментов Веба // Процессы управления и устойчивость. 2015. Т. 2(18). № 1. С. 458–463.

3. Blekanov I.S., Sergeev S.L., Maksimov A.I. Analysis of the topology of large Web segments using Broder's bow-tie model // Life Science Journal. 2014. Vol. 11. Pp. 258–261.

4. Ene A., Im S., Moseley B. Fast clustering using MapReduce // In KDD. 2011. Pp. 681–689.

5. Buhler J. Efficient large-scale sequence comparison by locality-sensitive hashing // Bioinformatics. 2001. No. 17(5). Pp. 419–428.

6. Broder A.Z. Identifying and Filtering Near-Duplicate Documents // Proc. of the 11th Annual Symp. on Combinatorial Pattern Matching. 2000. Pp. 1–10.

7. Jatsada Singthongchai, Suphakit Niwattanakul A Method for Measuring Keywords Similarity by Applying Jaccard's, N-Gram and Vector Space

// Lecture Notes on Information Theory. Vol. 1. No. 4. Pp. 159–164.

8. Chum O., Perdoch M., Matas J. Geometric minhashing: Finding a (thick) needle in a haystack // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2009. Pp. 17–24.

9. Gionis A., Indyk P., Motwani R. Similarity search in high dimensions via hashing // In VLDB. 1999. Pp. 518–529.

10. Koga H., Ishibashi T., Watanabe T. Fast agglomerative hierarchical clustering algorithm using Locality-Sensitive Hashing // Knowledge and Information Systems. 2007. Vol. 12. Iss. 1. Pp. 25–53.

11. Блеканов И.С., Сергеев С.Л., Мартыненко И.А. Построение тематико-ориентированных веб-краулеров с использованием обобщенного ядра // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. СПб.: Изд-во Политехн. ун-та, 2012. № 5(157). С. 9–15.

12. Marketing Search Engine and SEO Backlink Checker [электронный ресурс]/ URL: [www.majestic.com](http://www.majestic.com) (дата обращения: 02.04.2015).

## REFERENCES

1. *Ranking Web of Universities*. Available: [www.webometrics.info](http://www.webometrics.info) (Accessed: 10.11.2015).

2. Blekanov I.S., Moskalets R.Yu. Teoretiko-grafovyie kharakteristiki v vebometricheskikh issledovaniyakh vnutrenney topologii krupnykh segmentov Veba [Graph-theoretic characteristics webometric studies internal topology of large segments of the Web]. *Protsessy upravleniya i ustoychivost* [Stability and Control Processes], 2015, Vol. 2(18), No. 1, Pp. 458–463. (rus)

3. Blekanov I.S., Sergeev S.L., Maksimov A.I. Analysis of the topology of large Web segments using Broder's bow-tie model. *Life Science Journal*, 2014, Vol. 11, Pp. 258–261.

4. Ene A., Im S., Moseley B. Fast clustering using MapReduce. In *KDD*, 2011, Pp. 681–689.

5. Buhler J. Efficient large-scale sequence comparison by locality-sensitive hashing. *Bioinformatics*, 2001, No. 17(5), Pp. 419–428.

6. Broder A.Z. Identifying and Filtering Near-Duplicate Documents. In *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, 2000, Pp. 1–10.

7. Jatsada Singthongchai, Suphakit Niwattanakul A Method for Measuring Keywords Similarity by Applying Jaccard's, N-Gram and Vector Space. *Lecture Notes on Information Theory*, 2013, Vol. 1, No. 4, Pp. 159–164.

8. Chum O., Perdoch M., Matas J. Geometric minhashing: Finding a (thick) needle in a haystack. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009, Pp. 17–24.

9. Gionis A., Indyk P., Motwani R. Similarity search in high dimensions via hashing. In *VLDB*, 1999, Pp. 518–529.

10. Koga H., Ishibashi T., Watanabe T. Fast agglomerative hierarchical clustering algorithm using Locality-Sensitive Hashing. *Knowledge and Information Systems*, 2007, Vol. 12, Issue 1, Pp. 25–53.

11. Blekanov I.S., Sergeyev S.L., Martynenko I.A. Postroyeniye tematiko-oriyentirovannykh veb-kraulerov s ispolzovaniyem obobshchennogo yadra [The construction of the focused web-crawler using the universal kernel]. *Nauchno-tehnicheskie vedomosti SPbGPU. Informatika. Telekommunikatsii. Upravlenie* [St. Petersburg State Polytechnical University





*Journal. Computer Science. Telecommunications and Control System*], St. Petersburg: SPbGPU Publ., 2012, No. 5(157), Pp. 9–15. (rus)

12. *Marketing Search Engine and SEO Backlink Checker*. Available: [www. majestic.com](http://www.majestic.com) (Accessed: 02.04.2015).

---

**КОРЕЛИН Василий Николаевич** — аспирант кафедры технологии программирования Санкт-Петербургского государственного университета.

199034, Россия, Санкт-Петербург, Университетская наб., д. 7-9.

E-mail: [vn.korelin@gmail.com](mailto:vn.korelin@gmail.com)

**KORELIN Vasilii N.** *St. Petersburg State University.*

199034, Universitetskaya Emb. 7-9, St. Petersburg, Russia.

E-mail: [vn.korelin@gmail.com](mailto:vn.korelin@gmail.com)

**БЛЕКАНОВ Иван Станиславович** — доцент кафедры технологии программирования Санкт-Петербургского государственного университета, кандидат технических наук.

199034, Россия, Санкт-Петербург, Университетская наб., д. 7-9.

E-mail: [i.blekanov@gmail.com](mailto:i.blekanov@gmail.com)

**BLEKANOV Ivan S.** *St. Petersburg State University.*

199034, Universitetskaya Emb. 7-9, St. Petersburg, Russia.

E-mail: [i.blekanov@gmail.com](mailto:i.blekanov@gmail.com)

**СЕРГЕЕВ Сергей Львович** — заведующий кафедрой технологии программирования Санкт-Петербургского государственного университета, кандидат физико-математических наук, доцент.

199034, Россия, Санкт-Петербург, Университетская наб., д. 7-9.

E-mail: [slsergeev@yandex.ru](mailto:slsergeev@yandex.ru)

**SERGEEV Sergei L.** *St. Petersburg State University.*

199034, Universitetskaya Emb. 7-9, St. Petersburg, Russia.

E-mail: [slsergeev@yandex.ru](mailto:slsergeev@yandex.ru)